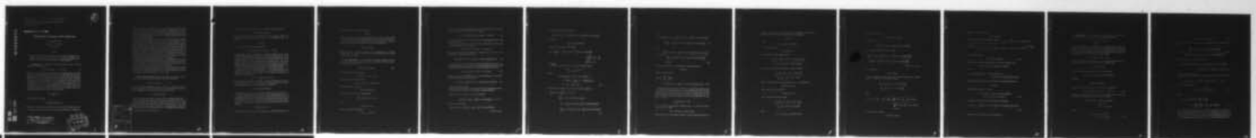
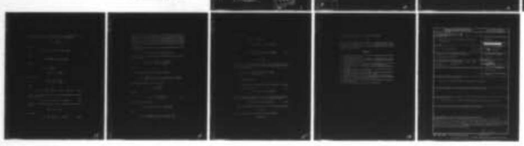
AD-A043 850 CALIFORNIA UNIV LOS ANGELES DEPT OF SYSTEM SCIENCE F/G 9/3
LIKELIHOOD RATIOS FOR SIGNALS IN ADDITIVE WHITE NOISE,(U)
JUN 77 A V BALAKRISHNAN AF-AFOSR-2492-73
UNCLASSIFIED AFOSR-TR-77-1150 NL

| OF |
AD
A043850





END
DATE
FILMED
9 -77
DDC

2

B.S.

ADA043850

AFOSR-TR- 77-1150

Likelihood Ratios for Signals in Additive White Noise*

A. V. BALAKRISHNAN

Department of System Science

UCLA, Los Angeles

ABSTRACT

We present a formula for likelihood functionals for signals in which the corrupting noise is modelled as white noise rather than the usual Wiener process. The main difference is the appearance of an additional term corresponding to the conditional mean square error. By way of one application we consider the 'order-disorder' problem of Shiryayev.

1. Introduction. The problem of detecting signals in noise is usually phrased in the following way:

$$y(t) = S(t) + N(t) \qquad 0 < t < T < \infty \tag{1}$$

where $y(\cdot)$ is the "observed" stochastic process being the sum of a "signal" process $S(t)$, and $N(t)$ the "noise" process, the two processes being mutually independent. The signal is usually a relatively "smooth" process in comparison with the noise—more specifically the noise process has "large" bandwidth compared to that of the signal (we use this bandwidth notion in a general way since the signal is not assumed to be necessarily stationary). To allow for this in a theoretical way, it was customary in engineering until the 1960's to allow $N(t)$ to be "white noise" even though unfortunately there was no precise definition of this. Because of this, beginning about 1960, it became fashionable to replace the model (1) by the "Wiener Process" version: We write

$$Y(t) = \int_0^t y(s)\,ds$$

and then make the change

$$Y(t) = \int_0^t S(\sigma)\,d\sigma + W(t),$$

where $W(t)$ is a Wiener process. There is no harm in this as long as operations

*Research supported in part under Grant No. 73-2492, Applied Math. Divn., AFOSR, USAF.

on the observed process are linear — as for example in the case of filtering theory for Gaussian signals. Significant difficulties appear in application as soon as we go to non-linear operations: as for example in calculating likelihood ratios (for continuous-time-parameter processes), basically because any non-linear operation is an Ito-integral in the observed data in which the differential dY cannot simply be replaced by $y(t)dt$.

Here we shall show that there is an alternate theory in which this difficulty does not arise. In this theory we go back to (1) but now take $N(t)$ to be "white-Gaussian" in the precise sense of inducing a weak distribution (finitely additive measure) on an L_2-space. In this case the process $y(\cdot)$ also induces a similar weak distribution and we may then calculate the Radon-Nikodym derivative of the weak distributions. This approach is many ways simpler; it avoids for example the necessity to invoke the Girsanov theorem. But the most important consequence is a calculable likelihood function. In this way we can rigorously justify the somewhat ad-hoc usage of the Wong-Zakai corrections or the Stratanovich integral or the circle differential of Ito. It turns out that the latter are special cases of the more general white noise theory.

In Section 2 we indicate the basic notions and definitions that we need. The main result—the likelihood functional formula is given in Section 3. In particular we see that there is a correction term to the usual Ito formula which can be calculated in terms of the conditional mean square error.

In Section 4 we indicate one application of the theory to the so-called "disruption"[†] problem initiated by Shiryayev-Kolmogorov [1]. We show that our formula for the conditional probability agrees with the version that has been used by Liptser [2] in actual analogue computer usage.

2. Basic Notions: White Noise Theory. We shall now indicate briefly the relevant notions from white noise theory, leaving details to [3]. Let

$$W(t) = L_2[(0,t); R_n], \qquad 0 < t < \infty$$

where R_n denotes real Euclidean space of dimension n. By white (Gaussian) noise we mean the elements of $W(t)$ under the weak distribution (Gauss measure) defined in terms of the characteristic function $C(h)$;

$$C(h) = \exp - \tfrac{1}{2}[h,h]_t$$

where $[\ ,\ ]_t$ denotes inner-product in $W(t)$. See for example Skorokhod [4].

More generally let μ denote any weak distribution on $W(t)$. Let $f(\cdot)$ be an function mapping $W(t)$ into another Hilbert space H_r. Let $f(\cdot)$ be Borel measurable. Let P be any finite dimensional projection on $W(t)$. Then $f(P\omega)$, where ω denotes points in $W(t)$, is referred to as a "tame function", since $f(P\omega)$ is a random variable in the sense that it induces a countably additive probability

[†]Also referred to as "order-disorder" problem.

measure on H_r so that in particular we may speak of the probability that

$$f(P\omega) \in \text{Borel set in } H_r.$$

We shall say that a Borel measurable function is a weak random variable if for any sequence of finite dimensional projections P_n, converging strongly to the identity:

$$\{ f(P_n\omega) \}$$

is a Cauchy sequence in probability such that

$$C(h) = \underset{n}{\operatorname{limit}} E\left[\exp i\left[f(P_n\omega), h\right]\right], \qquad h \in H_r$$

is independent of the particular sequence of projections chosen. In this case the weak distribution induced by $f(\omega)$ is by definition that corresponding to $C(h)$. If the space H_r is finite dimensional then of course the weak distribution will be countably additive so that we will drop the qualifier "weak". Also if the sequence $\{ f(P_n\omega)\}$ is Cauchy actually in the mean of order two, then by the Sazonov theorem we know that $C(h)$ must be the characteristic function of a countably additive measure and hence again we may drop the qualifier "weak" in this case as well. See [3] for many examples of such random variables.

Let μ_1, μ_2 be two weak distributions on $W(t)$ [or more generally on any separable Hilbert space]. We shall say that μ_2 is absolutely continuous with respect to μ_1 if given by $\epsilon > 0$ we can find $\delta > 0$ such that for any cylinder set C with finite-dimensional base:

$$\mu_2(C) < \epsilon \quad \text{whenever } \mu_1(C) < \delta.$$

We shall say that the random variable $f(\cdot)$ is the Radon-Nikodym derivative of μ_2 with respect to μ_1 if for any cylinder set C and any sequence P_n of finite dimensional projections converging strongly to the identity:

$$\mu_2(C) = \underset{n}{\operatorname{limit}} \int_C f(P_n\omega)\, d\mu_1$$

We shall be concerned with this only in the case where μ_1 is Gauss measure.

Finally a brief explanatory word about random variables (taking H_r now to be finite dimensional), and we take the case of Gauss measure. Let $\{\phi_i\}$ be any orthonormal basis in $W(t)$. Then we usually take the space of all sequences to be Ω, with β to be Borel sets therein and determine a countably additive measure on β which agrees with the given distributions of a finite number of the variables:

$$\int_0^t [\phi_i(s), \omega(s)]\, ds$$

In the case of Gauss measure, this is equivalent to taking Wiener measure on

$C[0,t]$. If $f(\cdot)$ is any random variable, then

$$\{ f(P_n\omega) \}$$

will be a Cauchy sequence converging in probability to some random variable measurable β (or its completion). With white noise however the sample space for us is l_2, the space of square summable sequences. In both cases we agree on distributions involving only a finite number of the basic variables:

$$\int_0^t [\phi_i(s), \omega(s)]\, ds$$

However not every variable measurable β can be expressed as a "random variable" $f(\omega)$ in our sense. For more on this see [3]. We shall not need to dwell on this much in this paper.

3. The Main Results. Let (Ω, β, p) denote a probability triple in the usual notation and $S(t,\omega)$ be an R_n valued jointly measurable stochastic process, $0 \leq t$, such that for each T, $0 < T$:

$$\int_0^T E\left[\|S(t,\omega)\|^2 \right] dt < \infty. \tag{3.1}$$

Since this implies in particular that

$$E\left[\|S(t,\omega)\|^2 \right] < \infty \quad \text{a.e. in } t,$$

we shall simply assume thruout that

$$E\left[\|S(t,\omega)\|^2 \right] < \infty \quad \text{for every } t, 0 < t.$$

Let $W(t)$ denote the L_2-space:

$$W(t) = L_2\left[(0,t]; R_n \right]$$

for each t, $0 < t$. Then by the Fubini theorem,

$$\int_0^T \|S(t,\omega)\|^2\, dt < \infty \quad \text{a.e. in } \omega$$

and

$$\phi(t;\omega) = S(\cdot,\omega),$$

where $S(\cdot,\omega)$ denotes the function

$$S(\sigma,\omega), \qquad 0 \leq \sigma \leq t$$

yields a (Borel) measurable mapping of Ω into $W(t)$, where we take the σ-algebra in $W(t)$ to be the Borel sets. In particular

$$C(t;h) = E\left[\exp i \int_0^t [S(\sigma,\omega), h(\sigma)]\, d\sigma\right] \qquad h \in W(t) \tag{3.2}$$

is, for each t, the characteristic function of a countably additive measure on the Borel sets of $W(t)$.

Next let $\eta(s,\omega')$, $0 \leq s \leq t < \infty$ denote white noise elements in $W(t)$, with sample points denoted by ω', and let

$$y(\sigma,\omega,\omega') = S(\sigma,\omega) + \eta(\sigma,\omega') \qquad 0 \leq \sigma \leq t < \infty.$$

We assume now [and thruout] that the processes $S(\cdot,\omega)$ and $\eta(\cdot,\omega')$ are independent and thus for $h(\cdot)$ in $W(t)$,

$$E\left[\exp i \int_0^t [y(\sigma,\omega,\omega'), h(\sigma)]\, d\sigma\right] = C(t;h) \exp - \frac{1}{2}\int_0^t \|h(\sigma)\|^2 d\sigma. \tag{3.3}$$

The first result we need is that the weak distribution induced by (3.3) is absolutely continuous with respect to that induced by the Gauss measure, and that the Radon-Nikodym derivative is a random variable $f(t,\cdot)$, given by:

$$f(t,h) = \int_{W(t)} \exp - \tfrac{1}{2}\{[S,S]_t - 2[S,h]_t\}\, d\mu_s, \tag{3.4}$$

where $[\ ,\ ]_t$ denotes inner product in $W(t)$, and μ_s the measure corresponding to (3.2). We note that we can rewrite this as:

$$f(t,h) = \int_\Omega \exp - \frac{1}{2}\left\{\int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), h(\sigma)]\, d\sigma\right\} dp. \tag{3.5}$$

This result is proved in [5].

We wish to show now that (3.5) for each h in $W(T)$ is absolutely continuous in $[0,T]$ with derivative given a.e. by (differentiating the integrand):

$$\frac{d}{dt} f(t,h) = \int_\Omega -\tfrac{1}{2}\{\|S(t,\omega)\|^2 - 2[S(t,\omega), h(t)]\}\exp$$
$$-\frac{1}{2}\left\{\int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), h(\sigma)\, d\sigma]\right\} dp \quad \text{a.e. } 0 < t < T.$$

First of all we note that

$$\int_\Omega \|S(t,\omega)\|^2 \exp - \frac{1}{2}\left\{\int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), h(\sigma)]\, d\sigma\right\} dp$$
$$\leq E[\|S(t,\omega)\|^2]$$

which is summable in $[0, T]$. Similarly,

$$\int_{\Omega}[S(t,\omega),h(t)]\exp-\frac{1}{2}\left\{\int_0^t\|S(\sigma,\omega)\|^2 d\sigma-2\int_0^t[S(\sigma,\omega),h(\sigma)]\,d\sigma\right\}dp$$

by Schwartz

$$\leqslant\sqrt{E[\|S(t,\omega)\|^2]\|h(t)\|^2}$$

which is also summable in $[0, T]$. Hence

$$\int_t^{t+\Delta}\int_{\Omega}\frac{d}{dt}\exp-\frac{1}{2}\left\{\int_0^t[S(\sigma),\omega]^2-2\int_0^t[S(\sigma,\omega),h(\sigma)]\,d\sigma\right\}dp$$

$$=\int_{\Omega}\left(\int_t^{t+\Delta}\frac{d}{dt}(\quad)\right)dp$$

$$=f(t+\Delta)-f(t)$$

as required.

Next let us note that, for any $\varepsilon>0$, we can find M such that

$$\int_{W(t)\cap(\|S\|\leqslant m)}d\mu_s\geqslant 1-\varepsilon.$$

Hence

$$\begin{aligned}
f(t,h)&\geqslant\left(\operatorname*{Inf}_{\|S\|\leqslant m}\exp-\tfrac{1}{2}\{\|S\|_t^2-2[S,h]_t\}\right)(1-\varepsilon)\\
&=(1-\varepsilon)\left(\exp-\tfrac{1}{2}\operatorname*{Sup}_{\|S\|\leqslant m}\{\|S-h\|_t^2\}\right)\left(\exp\tfrac{1}{2}\|h\|_t^2\right)\\
&\geqslant(1-\varepsilon)\exp-\tfrac{1}{2}(\|h\|_t+m)^2.
\end{aligned}\tag{3.6}$$

Hence (for $h(\cdot)$ in $W(T)$) we can take the logarithmic derivative:

$$\begin{aligned}
\frac{d}{dt}\log f(t,h)&=\frac{f'(t,h)}{f(t,h)}\quad\text{a.e. } 0\leqslant t\leqslant T\\
&=\left(\int_{\Omega}\{-\tfrac{1}{2}\|S(t,\omega)\|^2+[S(t,\omega),h(t)]\}\right.\\
&\quad\left.\times\exp-\frac{1}{2}\left\{\int_0^t\|S(\sigma,\omega)\|^2 d\sigma-2\int_0^t[S(\sigma,\omega),h(\sigma)]\,d\sigma\right\}dp\right)\Big/\\
&\quad\left(\int_{\Omega}\exp-\frac{1}{2}\left\{\int_0^t\|S(\sigma,\omega)\|^2 d\sigma-2\int_0^t[S(\sigma,\omega),h(\sigma)]\,d\sigma\right\}dp\right)
\end{aligned}\tag{3.7}$$

Let

$$\hat{S}(t;h) = \left(\int_{\Omega} S(t,\omega) \exp -\frac{1}{2} \left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), h(\sigma)] d\sigma \right\} dp \right) \Big/$$

$$\left(\int_{\Omega} \exp -\frac{1}{2} \left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), h(\sigma)] d\sigma \right\} dp \right) \quad (3.8)$$

Let

$$P(t,h) = \frac{\left(\int_{\Omega} \|S(t,\omega)\|^2 \exp -\frac{1}{2} \left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - \int_0^t [S(\sigma,\omega), h(\sigma)] d\sigma \right\} dp \right)}{\left(\int_{\Omega} \exp -\frac{1}{2} \left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - \int_0^t [S(\sigma,\omega), h(\sigma)] d\sigma \right\} dp \right)}$$

$$- \|\hat{S}(t,h)\|^2 \quad (3.9)$$

Then a simple application of the Schwartz inequality shows that

$$P(t,h) \geq 0.$$

Finally we have then that

$$\log f(t,h) = \int_0^t \frac{f'(s,h)}{f(s,h)} ds$$

$$= -\frac{1}{2} \left\{ \int_0^t \|\hat{S}(\sigma,h)\|^2 - 2\int_0^t [\hat{S}(\sigma,h), h(\sigma)] d\sigma + \int_0^t P(\sigma,h) d\sigma \right\}. \quad (3.10)$$

Substituting $y(\sigma,\omega,\omega')$ for $h(\cdot)$ in [3.10], we obtain the "log-likelihood function". We shall now show that $\hat{S}(t;y)$ has the interpretation as the conditional expectation of $S(t,\omega)$ given "$y(\sigma,\omega,\omega')$, $0 \leq \sigma \leq t$". The latter needs a more precise definition which we give now. Let $\{\phi_i\}$ be any orthornormal basis in $W(t)$ and let P_n correspond to the projection operator corresponding to the span of $\phi_1,\ldots,\phi_n$. Let

$$\zeta_i = \int_0^t \phi_i(\sigma) y(\sigma,\omega,\omega') d\sigma.$$

Let $\mathfrak{F}_n$ denote the sigma-algebra generated by $\zeta_1,\ldots,\zeta_n$. Let $P(t)y$ denote $y(\sigma,\omega,\omega')$, $0 \leq \sigma \leq t$. Then

$$E[S(t,\omega) | P_n \underline{P}(t) y] = E[S(t,\omega) | \mathfrak{F}_n]$$

yields a sequence of tame random variables, a Cauchy sequence in the mean of order two, being of course a martingale sequence (with finite second moment). We shall show that this sequence is equivalent to the sequence

$$\hat{S}(t; P_n y)$$

so that

$$\hat{S}(t; y) = E\left[S(t, \omega) | P(t) y \right].$$

For this first of all,

$$E\left[\zeta_1 | P_n \underline{P}(t) y \right]$$

is readily seen to be (Bayes Rule) given by:

$$\frac{\int_{W(t)} [S, \phi_1] \exp -\tfrac{1}{2} \{ \| P_n S \|_t^2 - 2 [P_n s, P_n y]_t \} \, d\mu_s}{\int_{W(t)} \exp -\tfrac{1}{2} \{ \| P_n S \|_t^2 - 2 [P_n s, P_n y]_t \} \, d\mu_s}$$

This is a Cauchy sequence in the mean of order two. Hence

$$\frac{\int_{W(t)} [S, \phi_1] \exp -\tfrac{1}{2} \{ \| S \|_t^2 - 2 [S, y]_t \} \, d\mu_s}{\int_{W(t)} \exp -\tfrac{1}{2} \{ \| S \|_t^2 - 2 [S, y]_t \} \, d\mu_s}$$

defines a random variable corresponding to the conditional expectation:

$$\operatorname*{limit}_n E\left[[S, \phi_1] | P_n y \right]$$

More generally, for any ϕ in $W(t)$,

$$\operatorname*{limit}_n E\left[[S, \phi] | P_n y \right]$$

is the identified with random variable:

$$\frac{\int_{W(t)} [S, \phi] \exp -\tfrac{1}{2} \{ \| S \|_t^2 - 2 [S, y]_t \} \, d\mu_s}{\int_{W(t)} \exp -\tfrac{1}{2} \{ \| S \|_t^2 - 2 [S, y]_t \} \, d\mu_s}$$

Now

$$E \| S(t, \omega) - n \int_{t - 1/n}^{n} S(\sigma, \omega) \, d\sigma \|^2 \to 0 \quad \text{a.e.}$$

8

and using the fact that

$$E(E[\zeta|P_n y])^2 \leqslant E[\zeta^2]$$

it follows that

$$\underset{n}{\text{limit}}\, E[\, S(t,\omega)|P_n y\,] = \lim_{m,n} E\left[n\int_{t-1/n}^{t} S(\sigma,\omega)\, d\sigma | P_m y \right]$$

can be identified with the random variable

$$\frac{\left(\int_{\Omega} S(t,\omega)\exp - \frac{1}{2}\left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [\, S(\sigma,\omega), y(\sigma)\,]\, d\sigma \right\} dp \right)}{\left(\int_{\Omega} \exp - \frac{1}{2}\left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [\, S(\sigma,\omega), y(\sigma)\,]\, d\sigma \right\} dp \right)}$$

$$= \hat{S}(t,y) \quad \text{a.e.} \qquad (3.11)$$

In a similar way

$$E\left[\, [\, S,\phi\,]_t^2 | P_n y \,\right]$$

is also a martingale convergent in the mean of order one, and we can see that the limit is identified with the random variable

$$\frac{\int_{W(t)} [\, S,\phi\,]_t^2 \exp - \frac{1}{2}\{\, \|S\|_t^2 - 2[\, S,y\,]_t \}\, d\mu_s}{\int_{W(t)} \exp - \frac{1}{2}\{\, \|S\|_t^2 - 2[\, S,y\,]_t \}\, d\mu_s}.$$

Again

$$E\left|\, \|S(t,\omega)\|^2 - \left\| \int_{t-1/n}^{t} nS(\sigma,\omega)\, d\sigma \right\|^2 \right| \leqslant \sqrt{E\,\|S(t,\omega) - \int_{t-1/n}^{t} nS(\sigma,\omega)\, d\sigma\|^2}$$

$$\sqrt{E\left[\, \|S(t,\omega) + \int_{t-1/n}^{n} nS(\sigma,\omega)\, d\sigma\|^2 \right]}$$

$$\to 0 \quad \text{with } n.$$

Also by Jensen's inequality:

$$E|E[\zeta|P_n]| \leqslant E[|\zeta|].$$

Hence we can readily infer that:

$$E\left[\,\|S(t,\omega)\|^2|P_n y\,\right]$$

converges in the mean of order one, and that

$$\frac{\int_{\Omega} \|S(t,\omega)\|^2 \exp -\frac{1}{2}\left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), y(\sigma)]d\sigma \right\} dp}{\int_{\Omega} \exp -\frac{1}{2}\left\{ \int_0^t \|S(\sigma,\omega)\|^2 d\sigma - 2\int_0^t [S(\sigma,\omega), y(\sigma)]d\sigma \right\} dp}$$

is a random variable

$$= E\left[\,\|S(t,\omega)\|^2|\underline{P}(t)y\,\right]. \tag{3.12}$$

Moreover we note that substituting $y(\sigma,\omega,\omega')$ for $h(\sigma)$ in [3.9], we have:

$$P(t;y) = E\left[\,\|S(t,\omega)\|^2|\underline{P}(t)y\,\right] - \|\hat{S}(t,y)\|^2.$$

Now

$$E\left[\,\|S(t,\omega)\|^2|P_n P(t)y\,\right] - \|E\left[\,S(t,\omega)|P_n\underline{P}(t)y\,\right]\|^2$$
$$= E\left[\,\|S(t,\omega) - E\left[\,S(t,\omega)|P_n\underline{P}_t y\,\right]\|^2|P_n\underline{P}_t y\,\right]$$

yielding a Cauchy sequence which is equivalent to the Cauchy sequence

$$P(t;P_n y).$$

Hence we can express $P(t,y)$ also as:

$$\begin{aligned} P(t,y) &= E\left[\,\|S(t,\omega) - E\left[\,S(t,\omega)|P(t)y\,\right]\|^2\underline{P}(t)y\,\right] \\ &= E\left[\,\|S(t,\omega) - \hat{S}(t;y)\|^2|\underline{P}(t)y\,\right]. \end{aligned} \tag{3.13}$$

Finally then we have for the log-likelihood-ratio functional:

$$\log f(t,y) = -\frac{1}{2}\left\{ \int_0^t \|\hat{S}(\sigma,y)\|^2 d\sigma - 2\int_0^t [\,S(\sigma,y),y(\sigma)\,]d\sigma + \int_0^t P(\sigma,y)d\sigma \right\}, \tag{3.14}$$

which is then a generalization of the formula in [5].

4. Applications. The first application of (3.14) is to the case where $S(t,\omega)$ is a Gaussian process because of the resulting simplification that

$$\begin{aligned} P(t,y) &= E\|S(t,\omega) - \hat{S}(t;y)\|^2 \\ &= P(t) \end{aligned}$$

and is thus independent of $y(\cdot)$. Moreover, it is possible, if $S(t,\omega)$ is further an Ito process, to obtain a differential characterization for $P(t)$ as well as of course for $\hat{S}(t;y)$. See [5]. In this paper, however, we shall concentrate on a non-Gaussian case, studied by Wonham [6] and Liptser-Shiryayev [7]. This is the case where $S(t,\omega)$ is a (one-dimensional) real-valued process:

$$S(t,\omega) = \sigma\chi(t-\tau(\omega)) \qquad t \geq 0,$$

where $\tau(\omega)$ is a Markov time relative to some growing sigma-algebra $B(t)$, and $\chi(t)$ is the characteristic function of the positive real-line, $t \geq 0$:

$$\chi(t) = \begin{cases} 1 & t \geq 0, \\ 0 & t < 0, \end{cases}$$

and σ is a fixed positive number. It is readily verified that:

$$E\left[S(t,\omega)^2\right] = \sigma E\left[S(t,\omega)\right] = \sigma^2 Pr\left[\tau(\omega) \geq t\right]$$

and hence

$$\int_0^T E\left(S(t,\omega)^2\right)dt < \infty.$$

The process $S(t,\omega)$ is of course jointly measurable in t and ω. Let

$$y(t,\omega,\omega') = \sigma\chi(t-\tau(\omega)) + \eta(t,\omega') \qquad 0 \leq t < \infty \tag{4.1}$$

where $\sigma > 0$, and $\eta(t,\omega')$ is white Gaussian noise in $W(T)$. Let us calculate the log-likelihood ratio based on $[0,T]$. First let us note that

$$\begin{aligned} E\left[S(t,\omega)^2 | \underline{P}(t)y\right] &= \sigma E\left[S(t,\omega) | \bar{P}(t)y\right] \\ &= \sigma\hat{S}(t;y) \\ &= \sigma^2 \qquad \pi(t,y) \geq 0 \end{aligned}$$

So that

$$\begin{aligned} P(t,y) &= \sigma\hat{S}(t;y) - \hat{S}(t;y)^2 \\ &= \sigma^2\left[\pi(t;y) - \pi(t;y)^2\right], \end{aligned} \tag{4.2}$$

where

$$\pi(t,y) = E\left[\chi(t-\tau(\omega))|P(t)y\right]. \tag{4.3}$$

Moreover we can readily calculate that, letting $P(\theta)$ denote the distribution of $\tau(\omega)$:

$$\pi(t,y) = \frac{\left(\int_0^t \exp -\frac{1}{2}\left\{\sigma^2(t-\theta) - 2\sigma\int_\theta^t y(s)\,ds\right\} dp(\theta)\right)}{\left(\int_t^\infty d\pi(\theta) + \int_0^t \exp -\frac{1}{2}\left\{\sigma^2(t-\theta) - 2\sigma\int_\theta^t y(s)\,ds\right\} dP(\theta)\right)}. \tag{4.4}$$

The log-likelihood ratio can thus be expressed:

$$-\frac{1}{2}\left\{\sigma^2\int_0^t \pi(\sigma,y)^2\,d\sigma - 2\sigma\int_0^t \pi(\sigma,y)y(\sigma)\,d\sigma + \sigma^2\int_0^t\left(\pi(\sigma,y) - \pi(\sigma,y)^2\right)d\sigma\right\}$$

$$= -\frac{1}{2}\left\{\sigma^2\int_0^t \pi(\sigma,y)\,d\sigma - 2\sigma\int_0^t \pi(\sigma,y)y(\sigma)\,d\sigma\right\}, \tag{4.5}$$

where $\pi(t,y)$ is given by (4.4).

The simplest case [going back to Wonham [6]] is to take for the distribution of $\tau(\omega)$:

$$Pr\left(\tau(\omega)=0\right] = P, \qquad Pr\left(\tau(\omega)>t>0\right] = (1-P)e^{-\lambda t}, \qquad \lambda>0.$$

In this case (4.4) becomes:

$$\pi(t,y) = \frac{N(t,y)}{(1-p)e^{-\lambda t} + N(t,y)}, \tag{4.6}$$

where

$$N(t,y) = \left\{\exp\left(-\frac{t\sigma^2}{2} + \sigma\int_0^t y(s)\,ds\right)\right\}p$$

$$+ (1-p)\lambda\int_0^t \exp\left\{\left(\frac{\sigma^2}{2} - \lambda\right)\theta - \sigma\int_0^\theta y(s)\,ds\right\}d\theta. \tag{4.7}$$

This is the same formula, except for replacement of $y\,dt$ by dY, given by Shiryayev [1]. But the difference appears more strikingly in the "differential equation" characterization of $\pi(t,y)$ between our version and the "Winer process" version. In our case we can derive an ordinary differential equation for $\pi(t,y)$ by differentiating directly with respect to t in (4.6), remembering that $y(\cdot)$

is an L_2-function over any finite interval. This is done most expeditiously in the following way: (similar to the technique in Shiryayev [1]). Rewrite:

$$\pi(t,y) = \frac{\phi(t)}{1+\phi(t)},$$

where

$$\phi(t) = \frac{1}{1-p}\, g(t)^{-1}\left(p + (1-p)\lambda \int_0^t g(s)\,ds\right),$$

where

$$g(t) = \exp\left(\left(\frac{\sigma^2}{2} - \lambda\right)t - \sigma \int_0^t y(s)\,ds\right).$$

Hence

$$\begin{aligned}\phi'(t) &= \lambda - \frac{g(t)'}{g(t)}\phi(t)\\ &= \lambda - \left(\frac{\sigma^2}{2} - \lambda - \sigma y(t)\right)\phi(t).\end{aligned}$$

Hence using

$$\frac{\pi'(t)}{\pi(t)} = \frac{\phi'(t)}{\phi(t)} - \frac{\phi'(t)}{1+\phi(t)}$$

we get

$$\pi'(t,y) = (1-\pi)\left(\lambda - \pi\frac{\sigma^2}{2}\right) + \pi(1-\pi)\sigma y(t); \qquad \pi(0,y) = p \tag{4.8}$$

which is then the white-noise version of the formula. In contrast to the Wiener process version in [1]:

$$d\pi = (1-\pi)(\lambda - \sigma^2\pi^2)\,dt + \sigma\pi(1-\pi)\,dY(t) \tag{4.9}$$

If we now add the Wong-Zakai [8] correction terms to this equation: namely subtract

$$\frac{\sigma^2}{2}(1-2\pi)(\pi-\pi^2)\,dt.$$

We obtain

$$d\pi = (1-\pi)\left(\lambda - \frac{\pi\sigma^2}{2}\right)dt + \pi(1-\pi)\sigma\,dY(t), \tag{4.10}$$

where we note that the only change is replacing $y(t)dt$ by $dY(t)$. Hence our solution is consistent with the Wong-Zakai correction, the Stratanovich integral [9] and the circle differential formalism of Ito [10]. Indeed in their calculations in actual simulation of $\pi(t,y)$, Dashevskii and Liptser [2] also appear to actually use [4.8] in place of the Ito-version (4.9) that they also derive. They obtain (4.8) in a purely formal manner however using the Stratanovich integral to replace the Ito integral.

As a slight (but nevertheless important in application) generalization, let us consider the case

$$y(t,\omega,\omega') = S_1(t) + \chi(t-\tau(\omega))S_2(t) + \eta(t,\omega'),$$

where $\chi(\ )$ and $\tau(\omega)$ are as before, and the distribution of $\tau(\omega)$ is the same. We can then calculate:

$$\hat{S}(t;y) = \frac{(1-p)S_1(t)e^{-\lambda t} + S_2(t)N(t)}{(1-p)e^{-\lambda t} + N(t)},$$

where

$$N(t) = p\exp -\frac{1}{2}\left\{\int_0^t S_2(\sigma)^2 d\sigma - 2\int_0^t S_2(\sigma)y(\sigma)d\sigma\right\}$$

$$+(1-p)\lambda\int_0^t \exp\left(-\frac{1}{2}\int_\theta^t S_2(\sigma)^2 d\sigma + \int_\theta^t S_2(\sigma)y(\sigma)d\sigma - \lambda\theta\right)d\theta.$$

If, following the previous example, we now take

$$\phi(t) = \frac{N(t)e^{\lambda t}}{(1-p)}$$

we have:

$$\hat{S}(t,y) = \frac{S_1(t) + S_2(t)\phi(t)}{1+\phi(t)}, \tag{4.11}$$

where $\phi(t)$ can be expressed as:

$$\phi(t) = \frac{1}{(1-p)}g(t)^{-1}\left[p + (1-p)\lambda\int_0^t g(s)ds\right],$$

where

$$g(t) = \exp\left(\frac{1}{2}\int_0^t S_2(\sigma)^2 d\sigma - \int_0^t S_2(\sigma)y(\sigma)d\sigma - \lambda t\right)$$

Hence, as before:

$$\phi'(t) = \lambda - \frac{g(t)'}{g(t)}\phi(t)$$

$$= \lambda - \left(\tfrac{1}{2}S_2(t)^2 - S_2(t)y(t) - \lambda\right)\phi(t)$$

or

$$\phi'(t) = \lambda - \left(\tfrac{1}{2}S_2(t)^2 - \lambda\right)\phi(t) + S_2(t)y(t)\phi(t) \qquad (4.12)$$

with

$$\phi(0) = \frac{P}{1-P}.$$

We can then also get a differential equation for $\hat{S}(t,y)$, assuming that $S_1(t)$ and $S_2(t)$ are differentiable, but we omit this detail here, excep to quote the result: (dot denoting derivative)

$$\dot{\hat{S}}(t,y) = \frac{(S_2(t) - \hat{S}(t,y))\dot{S}_1(t) + \dot{S}_2(t)(\hat{S}(t,y) - S_1(t))}{S_2(t) - S_1(t)}$$

$$+ \Big((S_2(t) + \hat{S}(t,y))$$

$$\times \left[\lambda(S_2(t) - \hat{S}(t,y)) - \left(\tfrac{1}{2}S_2(t)^2 - S_2(t)y(t) - \lambda\right)(\hat{S}(t,y) - S_1(t))\right]\Big) \Big/$$

$$(S_2(t) - S_1(t)) \qquad (4.13)$$

$$\hat{S}(0,y) = (1-p)S_1(0) + S_2(0)P.$$

Because of the rather complicated nature of this equation, perhaps it would be best to use (4.11) above with (4.12). The Ito version of (4.12) is given by:

$$d\phi = \lambda(1 + \phi(t))dt + S_2(t)\phi(t)dY(t). \qquad (4.14)$$

Finally we note that the statistics of $\hat{S}(t,y)$ are those of the Cauchy sequence

$$\hat{S}(t, P_n, y)$$

and in turn determined by (4.14) and (4.11). In particular then

$$[\,\mathrm{Inf}\ t, \hat{S}(t,y) > a\,]$$

is a Markov time with respect to the growing sigma algebra

$$\mathcal{B}(t)$$

where $\mathcal{B}(t)$ is the σ-algebra generated by $\{P_n v\}$; hence the statistics are determined by the "Wiener process" version. In particular then the optimal stopping time theory of Shiryayev for the "disorder" problem [1] can be exploited, mutatis mutandis.

References

[1] A. N. Shiryayev: Statistical Sequential Analysis, *A.M.S. Translations of Mathematical Monograph*, Vol. 38, Providence, 1973.

[2] M. L. Dashevskii and R. Sh. Liptser: Simulation of Stochastic Differential Equations Connected with the "Disorder" problem by Means of Analogue Computers. *Autmatikia i Telemekhanika* Vol. **27**, No. 4, 1966.

[3] A. V. Balakrishnan: A White Noise Version of the Girsanov Formula: *Proceedings of the Symposium on Stochastic Differential Equations*, Kyoto 1976, edited by K. Ito.

[4] A. V. Skorokhod: Integration in Hilbert Space, Springer-Verlag, Berlin, Heidelberg, New York, 1974.

[5] A. V. Balakrishnan: Radon Nikodym Derivatives of a Class of Weak Distributions on Hilbert Spaces. *Appl. Math. Opt.* **3** (1977) 209–225.

[6] W. M. Wonham: Some Applications of Stochastic Differential Equations to Optimal Non-linear Filtering *J. SIAM on Control, Ser. A*, Vol. **2**, No. 3, 1965.

[7] R. S. Liptser and A. N. Shiryayev: *Statistics of Random Processes*, Nauka 1975 (Russian).

[8] E. Wong and M. Zakai: On the Relation Between Ordinary Integrals and Stochastic Integrals, *Intern. J. of Engrg. Science*, **3** (1965) p. 213–229.

[9] R. L. Stratanovich: *Conditioned Markov Processes and Their Application to Optimal Control*, Elsevier, New York 1968.

[10] K. Ito: Stochastic Differentials, *Appl. Math. Opt.* **1**, (1976), 374–381.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 19

12 AFOSR-TR-77-1150

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

6 LIKELIHOOD RATIOS FOR SIGNALS IN ADDITIVE WHITE NOISE,

5. TYPE OF REPORT & PERIOD COVERED

Interim

6. PERFORMING ORG. REPORT NUMBER

[redacted]

7. AUTHOR(s)

10 A. V. Balakrishnan

8. CONTRACT OR GRANT NUMBER(s)

15 AF-AFOSR-~~73~~-2492-73

9. PERFORMING ORGANIZATION NAME AND ADDRESS

University of California, Los Angeles,
System Science Department
Los Angeles, CA 90024

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

61102F
16 2304/A1 17 A1

11. CONTROLLING OFFICE NAME AND ADDRESS

Air Force Office of Scientific Research (NM)
Building 410 Bolling AFB
Washington, D. C. 20332

12. REPORT DATE

11 June 1977

13. NUMBER OF PAGES

16 12 17p.

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

To be published in Journal of Math Analysis and Applications, 1977

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

is presented

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

~~We present~~ a formula for likelihood functionals for signals in which the corrupting noise is modelled as white noise rather than the usual Wiener process. The main difference is the appearance of an additional term corresponding to the conditional mean square error. ~~By way of~~ one application is to ~~we~~ consider the 'order-disorder' problem of Shiryayev.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

465213

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)